



UNITED STATES PATENT AND TRADEMARK OFFICE

A

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/762,126	02/02/2001	Shinji Miwa	P5276b	1392

20178 7590 07/12/2005

EPSON RESEARCH AND DEVELOPMENT INC
INTELLECTUAL PROPERTY DEPT
150 RIVER OAKS PARKWAY, SUITE 225
SAN JOSE, CA 95134

EXAMINER

BLACKWELL, JAMES H

ART UNIT	PAPER NUMBER
----------	--------------

2176

DATE MAILED: 07/12/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/762,126

Applicant(s)

MIWA ET AL.

Examiner

James H. Blackwell

Art Unit

2176

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 25 April 2005.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-5 and 7-23 is/are pending in the application.
- 4a) Of the above claim(s) 6 is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☐ Claim(s) 1-5 and 7-23 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 02 February 2001 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some * c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: _____

RD

DETAILED ACTION

1. This Office Action is in response to Response B received on 04/25/05.
2. Claims 1-5, and 7-23 are pending.
3. Claims 1 and 7 are currently amended.
4. The rejection of independent Claims 1 and 7 (and their corresponding dependent claims) under 35 U.S.C. 101 have been withdrawn as necessitated by amendment. Similar rejections with respect to Claims 10, and 11 have been withdrawn. However, the Examiner now rejects Claims 10, and 11 under 35 U.S.C. 101 for the reasons stated below.

Claim Rejections - 35 USC § 101

5. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Independent apparatus claims 10, and 11 and dependent apparatus claims 14-17, and 20-21 appear to be drawn to a non-tangible, software arrangement per se. The elements of the claims appear to be either software programs or combinations of software programs and data with no hardware implementation required, and are therefore non-statutory under 35 U.S.C. 101.

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 1-5, 10, 12, 14, and 16-19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Zamir et al. (hereinafter Zamir, "Web Document Clustering: A Feasibility Demonstration", ACM, August 1998).

In regard to independent Claim 1 (and similarly independent Claims 10, and 12), Zamir teaches *a document categorizing method for categorizing a plurality of documents into a plurality of clusters according to semantic similarity* in that the STC algorithm, which is a linear time clustering algorithm. STC has three logical steps: (1) document cleaning, (2) identifying *base clusters* using a suffix tree, and (3) merging the *base clusters into clusters* (p. 48, Col. 1, Sec. 3, lines 18-25).

Zamir also teaches *a cluster merging process is performed such that relations among clusters of said plurality of clusters are evaluated on the basis of documents included in the respective clusters* in that step (2) of the STC algorithm, the identification of base clusters can be viewed as the creation of an inverted index of phrases for our document collection. This is done efficiently using a data structure called a *suffix tree*. This structure can be constructed in time linear with the size of the collection, and can be constructed incrementally as the documents are being read (p. 48, Col. 1, Sec 3.2,

lines 43-49). Each base cluster is assigned a score that is a function of the number of documents it contains, and the number of words that make up its phrase (p. 48, Col. 2, Sec 3.2, lines 30-32).

Zamir also teaches *two or more clusters having a degree of relation equal to or higher than a predetermined value are combined together* in that the final step of the STC algorithm merges base clusters with a high degree of overlap in their document sets (p. 49, Col. 1, lines 19-21).

Zamir fails to teach that *said cluster merging process defines said degree of relation between multiple clusters under consideration as the number of distinct files common to all of said clusters under consideration multiplied by a predefined multiplication factor divided by a total sum of all the files in said clusters under consideration*. However, since what is claimed is simply a variation of Dice's coefficient, one of many similarity measures that are commonly known in the art, it would have been obvious to one of ordinary skill in the art at the time of invention to use any one of the possible similarity measures to assist in determining whether or not two clusters should be combined.

In regard to dependent Claim 2 (and similarly dependent Claims 16 and 18), Zamir fails to specifically teach that *said multiplication factor is equal to the number of clusters under consideration*. However, since what is claimed is simply a variation of Dice's coefficient, one of many similarity measures that are commonly known in the art, and that when using Dice's coefficient for comparing two clusters, the multiplication

factor is normally equal to the number of clusters under consideration, it would have been obvious to one of ordinary skill in the art at the time of invention to use any one of the possible similarity measures to assist in determining whether or not two clusters should be combined.

In regard to dependent Claim 3, Zamir teaches that *said cluster merging process is performed such that the manner in which feature elements, which characterize respective clusters under consideration as to whether they should be merged or not, appear in the respective clusters under consideration is examined, and cluster merging is performed in accordance with the manner in which the feature elements appear in that each base cluster is assigned a score that is a function of the number of documents it contains, and the number of words that make up its phrase (p. 48, Col. 2, Sec 3.2, lines 30-32).*

In regard to dependent Claim 4, Zamir teaches *said cluster merging process is performed at least for two clusters, and after completion of the cluster merging process a first time, said cluster merging process is repeatedly performed on the resultant set of clusters until no further cluster merging occurs* in that in essence, we are clustering the base clusters using the equivalent of a single-link clustering algorithm where a predetermined minimal similarity between base clusters serves as the halting criterion (implying that it keeps clustering clusters until a condition is met) (p. 49, Col. 1, Sec 3.3, lines 40-41; Col. 2, lines 1-2).

In regard to dependent Claim 5, Zamir teaches *after completion of said cluster merging process, supplementary information indicating that cluster merging has been*

performed and also indicating the basis on which the cluster merging has been performed is output in that Fig. 1 output of the clustering process (p. 47).

In regard to dependent Claim 14 (and similarly dependent Claims 17, and 19), Zamir fails to teach that *said multiplication factor and said number of clusters under consideration is two*. However, since what is claimed is simply a variation of Dice's coefficient, one of many similarity measures that are commonly known in the art, and that when using Dice's coefficient for comparing two clusters, the multiplication factor is normally equal to the number of clusters under consideration, it would have been obvious to one of ordinary skill in the art at the time of invention to use any one of the possible similarity measures to assist in determining whether or not two clusters should be combined.

7. Claims 7-9, 11, 13, 15, and 20-23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Zamir in view of Wu (U.S. Patent No. 5,991,756).

In regard to independent Claim 7 (and similarly independent Claims 11, and 13), Claim 7 (and similarly Claims 11, and 13) reflects the document categorizing method as Claimed in Claim 1, and is rejected along the same rationale.

In addition, Zamir fails to specifically teach that ... *said cluster names are displayed in a first listing format, and when said degree of relation among said clusters is lower than said second predetermined value and higher than said first predetermined value, said cluster names are displayed in a second listing format*. However, Wu teaches in Fig. 5 the display of a Yahoo search result that might result from submitting

the query string "The game of go" to their search engine. Listed are a series of category names (*cluster names*) listed in a hierarchical format, which are links to groups of similar documents (*clusters*). Though Wu does not call these categories/sub-categories names clusters, the fact that each link in the hierarchy from left to right (and from top to bottom) represents a group of similar documents, by definition can be thought of as clusters of similar documents. As one traverses the hierarchy from left to right, one traverses the cluster hierarchy from general to more specific. This traversal also inherently represents a degree of similarity of documents.

This method of displaying hierarchical/related structures is often referred to as a "Bread Crumb Trail". Additionally, one could also display such a structure through the use of a dendrogram or tree (in the case of a vertical display, similar to Zamir, Fig. 2).

Though not specifically taught by Wu, it would have been obvious to one of ordinary skill in the art at the time of invention to conclude that such a portrayal of document cluster names as seen in Figure 5 constitutes the claimed first and second listing formats based on interpretation of similarity measures (Col. 8, lines 46-56). It would have been obvious to one of ordinary skill in the art at the time of invention to combine the teachings of Zamir and Wu as both inventions relate to grouping documents based on their similarities. The addition of Wu provides the benefit of a method of presenting the document hierarchies as a function of similarity that is easy to understand.

In regard to dependent Claim 8, Zamir fails to specifically teach that *when said cluster names are displayed in said first listing format, said cluster names of the*

respective clusters are displayed successively in a single horizontal line or are displayed successively in different lines. However, Wu teaches in Figure 5 a hierarchy of document clusters (see argument in Claim 7) that are listed in a single line (54, 56, 58) as well as being displayed on different lines.

Zamir also fails to teach that *when said cluster names are displayed in said second listing format, a delimiter is inserted between adjacent cluster names of the respective clusters.* However, Wu teaches in Fig. 5 listings of clusters separated by a colon delimiter (54, 56, 58). It would have been obvious to one of ordinary skill in the art at the time of invention to combine the teachings of Zamir and Wu as both inventions relate to grouping documents based on their similarities. The addition of Wu provides the benefit of a method of presenting the document hierarchies as a function of similarity that is easy to understand.

In regard to dependent Claim 9, Zamir fails to teach that *when a first cluster includes a second cluster therein, the name of said second cluster included in said first cluster is enclosed within brackets and placed after the name of said first cluster.* However, Wu teaches in Fig. 5 listings of clusters separated by a colon delimiter (54, 56, 58). Though not delimiting by brackets as claimed, it would have been obvious to one of ordinary skill in the art at the time of invention to combine the teachings of Zamir and Wu as both inventions relate to grouping documents based on their similarities. The addition of Wu provides the benefit of a method of presenting the document hierarchies as a function of similarity that is easy to understand.

Art Unit: 2176

In regard to dependent Claim 15 (and similarly dependent Claims 20, and 22), Claim 15 (and similarly Claims 20, and 22) teach methods for categorizing documents as taught in Claim 7 (and similarly Claims 11, and 13) and are rejected along the same rationale.

In regard to dependent Claim 21 (and similarly dependent Claim 23), Claim 21 (and similarly Claim 23) teach methods for categorizing documents as taught in Claim 8, and are rejected along the same rationale.

Response to Arguments

8. Applicant's arguments with respect to claims 1-5, 10, 12, 14, and 16-18 have been considered but are moot in view of the new ground(s) of rejection. Specifically, the claimed degree of relation reads on nothing more than perhaps a variation of Dice's similarity measure, which along with numerous other measures of similarity are used in clustering techniques to determine whether or not two clusters (in the case of an agglomerative or "bottom up" clustering method) belong together in the same cluster. Agglomerative clustering, by definition, assumes that each document forms an initial cluster and from there on, clusters are combined based on a similarity measure. Though the specific similarity measure is not taught by Zamir, it would have been just as obvious to use any measure of similarity in determining whether or not to cluster two clusters together.

9. Arguments are made with respect to Claims 7, 11, 13 in the application of the prior art of Wu (U.S. Patent No. 5,991,756) by the applicant wherein the prior art of Zamir in view of Wu fails to teach the limitation *wherein the cluster names of respective clusters merged together are displayed such that when said degree of relation among said clusters is higher than a second predetermined value higher than said first predetermined value, said cluster names are displayed in a first listing format, and when said degree of relation among said clusters is lower than said second predetermined value and higher than said first predetermined value, said cluster names are displayed*

Art Unit: 2176

in a second listing format. The Examiner disagrees, since by admission, the Applicant states that Wu teaches one naming format (p. 10, Response B), but not two naming formats as claimed, and that the decision on which of the two naming formats is used depends upon the results of similarity results. The Examiner would argue that both Zamir and Wu teach different display methods for the portrayal of cluster names (see revised rejection above), so that in combination they teach two different display methods. As for the argument describing that, which display is chosen depends on a value of a "similarity measure" between cluster names, the examiner respectfully points out that Zamir teaches, by way of Fig. 2, in a hierarchical fashion, such a relationship between a similarity measure and the cluster names (their location in the tree depicts how similar or not one cluster is from another).

Conclusion

10. Any inquiry concerning this communication or earlier communications from the examiner should be directed to James H. Blackwell whose telephone number is 571-272-4089. The examiner can normally be reached on Mon-Fri.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather R. Herndon can be reached on 571-272-4136. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306 (after July 15th, the new number will be 571-273-8300).

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

James H. Blackwell
07/03/05

William L. Bashore
WILLIAM BASHORE
PRIMARY EXAMINER
7/10/2005